PTO 03-1105 HAMT

Japanese Patent
Document No. 09-075085

translation

A constituent of human 26S proteasome and human DNA that codes the constituent

[ヒト２６Ｓプロテアソーム構成成分蛋白質およびそれをコードするＤＮＡ]

Kato Seishi

UNITED STATES PATENT AND TRADEMARK OFFICE
Washington, D.C. 12/2002

Translated by: Schreiber Translations, Inc.

Bibliographic Fields
Document Identity
(19) [Publication Office]
Japan Patent Office (JP)
(12) [Kind of Document]
Publication of Unexamined Patent Application (A)
(11) [Publication Number of Unexamined Patent Application]

Japan Unexamined Patent Publication Hei 9 - 75085

(43) [Date of Publication of Unexamined Patent Application]

25th March 1997 102 (b)

(43) [Date of Publication of Unexamined Patent Application]

25th March 1997

(54) [Title of Invention]
A constituent of human 26S proteasome and human DNA that codes the constituent

(51) [International Patent Classification, 6th Edition]

C12N 15/09 ZNA

C07H 21/04

C12N 1/21

Sep-64

//A61K 38/00 AAM

ADU

C12N 15/09 ZNA

C12R 1:91)

C12N 1/21

C12R 1:19)

C12N 9/64

C12R 1:19)

[FI]

C12N 15/00 ZNA A 916 2- 4B

C07H 21/04 B

C12N 1/21

Sep-64

A61K 37/02 AAM

ADU

[Number of Claims]

4

[Form of Application]

OL

[Number of Pages in Document]

9

[Request for Examination]

Not yet requested

(21) [Application Number]

Japan Patent Application Hei 7 - 235052

(22) [Application Date]

1995 (1995) September 1 3 days

(71) [Applicant]

[Identification Number]

173762

[Name]

SAGAMI CHEMICAL RESEARCH CENTER (DB 69-172-6959)

[Address]

Kanagawa Prefecture Sagamihara City Nishi Onuma 4-Chome 4-1

(72) [Inventor]

[Name]

Kato Seishi

[Address]

Kanagawa Prefecture Sagamihara City Nandai 1 - 9 - 2

(72) [Inventor]

[Name]

Saeki -------

[Address]

Kanagawa Prefecture Chigasaki City Matsukazedai 17 - 20

(72) [Inventor]

[Name]

Sekine Shingo

[Address]

Kanagawa Prefecture Sagamihara City Nishi Onuma 4 - 4 - 1

(72) [Inventor]

[Name]

Tanaka Keiji

[Address]

Tokushima Prefecture Tokushima City Shimomachi down town, lane no 206

(57) [Abstract]

[Problems to be Solved by the Invention]

To provide a constituent of human 26S proteasome and human DNA that codes the constituent such that, the constituent plays an important role in intracellular protein degradation and can be used in diagnosis and treatment of various types of diseases.

[Means to Solve the Problems]

Protein, which includes amino acid sequence displayed in sequence number 1 and a DNA that codes said protein, for example a DNA that includes base sequence displayed in sequence number 1.

[Effect(s)]

Said protein is acquired by revealing recombinant of human DNA that codes human proteasome constituent P28 protein of this invention.

[Scope Of Claim(s)]

[Claim 1]

Protein that includes amino acid sequence displayed in sequence number 1.

[Claim 2]

DNA that codes the protein stated in claim 1.

[Claim 3]

DNA that includes the base sequence displayed in sequence number 1.

[Claim 4]

DNA stated in claim 3 comprises of the base sequence displayed in sequence number 2.

[Detail Description of the Invention]

[0001]

[Technological Field of Invention]

This invention pertains to protein, which constitutes 26S proteasome in human intracellular protease and to a DNA which codes that

The protein of this invention is not only useful in elucidating the functions of human 26S proteasome but also in diagnosing and treating various types of diseases.

The human DNA of this invention can be used as a probe for genetic diagnosis and as a gene resource for genetic therapy.

In addition, it can be used as a gene resource for the mass production of the protein that is being coded by said DNA.

[0002]

[Prior Art]

Proteasome, which is a multifunctional protease, is such an enzyme that exists widely in the eukaryote that reaches to human from yeast and degrades the energy dependent ubiquitination protein.

Proteasome is structured with 20S proteasome that is comprised of variety of constituents having molecular mass of 21 to 31kDa and with PA700 controlling proteins of 30 to 112kDa, having the sedimentation coefficient 22S, as a whole, it is structured with a macromolecule; (later called as 26S proteasome) which has a sedimentation coefficient 26, approximate molecular mass is 2 million Da. , [Rechchsteiner, et al., Journal of Biological Chemistry (0021 - 9258, JBCHA3). 268: 6065 - 6068 (1993), Yoshimura、T. et al.、J. St ruct. Bi ol.111: 200-211 (1993), Tanaka、K. et al., New Biologist 4 : 1 7 3- 187 (1992)].

The full capacity of proteasome has not become clear however the following functions and usefulness has become clear from the research done using yeast and mouse among others.

[0003]

In intracellular of eukaryote, energy (ATP) dependent protein degradation is started due to the fact that ubiquitine connects with protein selectively but the fact has become clear is that the protein degrading energy dependent activity is in proteasome, especially in 26S proteasome [Chu-Ping, M. etal, Journal of Biological Chemistry (0021 - 9258, JBCHA3). 269: 3539 - 3547 (1994)],

human 26S proteasome is considered to be useful in the clarification of the energy dependent protein degradation mechanism.

[0004]

It has been revealed that the degradation of cell cycle related genes such as oncogenes and cyclin including c-Myc, Mos, Fos is conducted by energy and ubiquitine dependent 26S proteasome [Ishida, N. et al., FEBS Letters (0014 - 5793, FEBLAL) 324: 345 - 348 (1993), Hershko, A. and Ciecha nover, A., Annu. rev. Biochem. 61:76 1- 807 (1992)], and the importance of the proteasome in the cell cycle control has been recognized.

[0005]

In addition, proteasome gene is developed abnormally in liver cancer cell, renal cancer cell and leukemia cell among others and [Kanayama, H. et al., Cancer Res.51: 6677-6685 (1991)], it is observed that proteasome is accumulated to abnormality in tumor cell nucleus.

Therefore, human proteasome is considered to be useful in the elucidation of the cancer mechanism and in the diagnosis or the treatment of cancer

[0006]

In addition, the development of proteasome is induced with interferon γ or the like and the fact is suggested that it is deeply involved in intracellular class I major histocompatible complex presentation. [Aki、 M. et al., Journal of Biochemistry (0021 - 924 X, JOBIAO) 115: 257 - 269 (1994), Micha lek, M.T. et al., Nature (London) (0028 - 0836) 363: 552 - 5554 (1994)].

Therefore, the constituent component of the human proteasome can be utilized in the explanation of the mechanism of antigen presentation of immune system and in the development of immune suppressing drug.

[0007]

Furthermore, from the phenomena that the ubiquitination protein is accumulated abnormally in the brains of Alzheimer patients, [Kitaguchi, N. et al., Nature (London) (0028 - 0836) 331: 530 - 532 (1988)] it is suggested that the proteasome is involved in Alzheimer disease and a human proteasome is considered to be useful in the clarification of the cause of Alzheimer and in it's treatment.

[0008]

It has been disclosed in Japan Unexamined Patent Publication Hei 5-292964 with regard to the protein that possesses the characteristic of human 26S proteasome, concerning rat proteasome constituent protein it is disclosed in Japan Unexamined Patent Publication Hei 5-268957 and in Japan Unexamined Patent Publication Hei 5-31 7059 however, regarding the human 26S proteasome constituent component of this invention is not known.

[0009]

[Problems to be Solved by the Invention]

The objective of this invention is to provide such a protein, the molecular weight of which is approximately 28 k Da (later called as P28 protein)and that which forms human 26S proteasome and a DNA that codes said protein.

[0010]

[Means to Solve the Problems]

As a result of diligent research, the inventors did cloning of human cDNA that codes P28 protein, which constitutes human 26S proteasome and completed this invention.

In other words, this invention provides a protein, which is a human P28 protein and that which includes amino acid sequence, which is displayed in sequence number 1.

In addition this invention provides a DNA that codes the above-described protein, e.g. cDNA, which includes base sequence that is displayed in sequence number 1.

[0011]

[Embodiment of the Invention]

The protein of this invention can be acquired by the method of isolating from human internal organ or cell line, by the method of preparing peptide with the chemical synthesis based on the amino acid sequence of this invention or by the production method with the DNA transfer technique using the DNA that codes human P28 protein of this invention but the method of acquiring with DNA transfer technique is preferred.

For example, RNA is prepared by transcribing in-vitro from vector that possesses cDNA of this invention; in-vitro can be developed by conducting in-vitro translation with this as a matrix.

In addition if the translation area is transferred to the suitable developed vector by with the known method, it is possible to develop the protein that codes with colon bacillus, hay bacillus, yeast, animal cell among others on the large scale.

[0012]

All DNA that code above-mentioned protein are included in DNA of this invention.

Said DNA can be acquired by using the methods such as chemical synthesis, cDNA cloning.

[0013]

The cDNA of this invention, for example it is possible to clone from cDNA library derived from human cell.

cDNA synthesizes by templating poly $(A)^+$RNA that is extracted from human cell.

As a human cell, depending on the surgeries on the human body, extraction is all right or even cultured cell is all right.

In the embodiment poly (A)+RNA isolated from human phosphorous cell U937 was used.

It is all right to synthesize cDNA by using Okayama-Berg method [Okayama, H. and Berg, P., Molecular and Cellular Biology (0270 - 7306, MCEBD4) 2: 161 - 170 (1982)], Gubler-Hoffman method [Gubler. and Hoffman, J. Gene (0378 - 1119, GENED6) 25: 263 - 269 (1983)], but in order to obtain full length clone efficiently, the use of vector primer as given in the embodiment is advisable.

[0014]

Cloning of cDNA is to be conducted with isolation refinement of P28 protein, which is a constituent component of bovine 26S proteasome and partial amino acid sequence determination, partial base sequence determination of cDNA clone that is arbitrarily selected from cDNA library, database creation of amino acid sequence that is predicted from the base sequence and the database search based on the partial amino acid sequence of bovine P28 protein.

Identification of cDNA is to be conducted with complete base sequence determination based on sequencing, with comparison of amino acid sequence predicted from the base sequence and bovine P28 protein partial amino acid sequence, with the protein development due to in-vitro translation and with the development due to colon bacillus.

[0015]

The cDNA of this invention is characterized by the fact that it includes the base sequence displayed in sequence number 1, for example, those displayed in sequence number 2 have base sequence that consists of 1468bp and open reading frame of 681bp.

This open reading frame codes the protein that consists of 226 amino acid residue.

[0016]

Furthermore, the clone same as cDNA of this invention can be easily acquired by using an oligonucleotide probe, which is synthesized based on the base sequence of cDNA that is stated in sequence number 1 and sequence number 2 and by screening human cDNA library that is produced from the cell line of this invention.

[0017]

Generally, human gene with multi types depending on the individual differences is recognized in frequent.

Therefore pertaining to Sequence Number 1 or Sequence Number 2, cDNA substituted due to addition of one or plurality of nucleotide, depending on the deletion and/or other nucleotide, also comes in the category of this invention.

[0018]

In the same way, occurring due to these modifications, protein substituted depending on the addition of one or plurality of amino acid and due to the deletion and/or other amino acid also comes in the category of this invention given that it possesses the activity of protein that has amino acid sequence displayed in sequence number 1.

[0019]

In the cDNA of this invention cDNA fragment (10bp or more) is included, which includes every partial base sequence of the base sequence displayed in sequence number 1 or 2

In addition, also DNA fragment that consists of sense chain and antisense comes into this category.

These DNA fragments can be used as probe for gene therapy.
[0020]
[Working Example(s)]
Next this invention is explained concretely with working embodiment, however, this invention does not limit itself to these examples.
Basic operation and enzyme reaction pertaining to DNA transfer was according to the literature.

["Molecular Cloning. A Laboratory Manual", Cold Spring Harbor Laboratory, 1989].
Restriction enzyme and various modified enzymes of Takara Shuzo Co. Ltd were used; especially when above stated was absent.
Buffer solution composition of each enzyme reaction and reaction conditions was according to the attached explanatory manual.

cDNA synthesis was according to the literature [Kato、S. et al.、 Gene (0378 - 1119, GENED6) 150: 243 - 250 (1994)].
[0021]
Isolation and purification of bovine 26Sproteasome constituent P28 protein and determination of partial amino acid sequence.

According to the bovine purification method described in literature [Chu-Ping、M. et al., Journal of Biological Chemistry (0021 - 9258, JBCHA3). 269: 3539 - 3547 (1994)], purification of bovine proteasome is conducted with column chromatography, which uses ammonium sulphate deposits, Sephacryl S-300, DEAE flacto gel and hydroxyapatite from bovine red blood cells..
From the acquired bovine proteasome P28 protein was fractionated with high performance liquid chromatography (HPLC).

Said elution fraction, was conducted under dithiothreitol reduction, 10% SD S-PAGE and bovine P28 protein was isolated and purified.

[0022]

Partial amino acid sequence of bovine P28 protein was determined with the method below.

Bovine P28 protein, which is separated by SD S-PAGE, in 0.1 Mtris buffer solution (pH 9.0), enzyme digestion was conducted for 8 hours in 4 M urea, with 1μg of lysine specific endoprotease at the temperature of 37 deg C.

Acquired partial peptide fragment was separated with reverse HPLC, regarding the 4 types of peptide fragments N terminal amino acid sequence was determined by automatic protease sequencer (Applied Biosystems Corporation).

N-terminal amino acid sequence of each peptide fragment was shown in Sequence Number 3~6.
[0023]
Poly (A)+RNA manufacturing
After culturing human phosphorus former cell line U937 (ATCCCR L 1593) in the culture medium of RPMI1640 which includes 5% fetal calf serum under 5% of CO_2 air current at 37 deg. C it was treated for 16 hours in phorbol myristate (30 ng/ml medium) and the cell of 1.1g was acquired.
After melting this in 5.5 M **guar Ni di ウム thiocyanate** solution 16 ml, mRNA was manufactured in accordance with literature [Okayama et al.、"Methods in Enzymology (0076 - 6879) "Vol.164、 Academic Press,1987].
This was washed with 20m Mtris-HCL (pH7.6), 0.5 M NaCL, 1m MEDTA and was kept in oligo dT cellulose column and then in accordance with the above mentioned literature poly (A) >+RNA was refined.
Poly of 72;mu g (A) +RNA was acquired this way.
[0024]
Creation of cDNA library
After cloning vector pkA 1 (Japan Unexamined Patent Publication Hei 4- 117292 disclosure) is digested with Kpnl, approximately 60 dT tail were added with terminal transferring enzyme.

This was used as vector primer by EcoRV digestion and by removing dT tail of the one side.

The reaction conditions of cDNA synthesis were in accordance with the literature [above stated literature on Okayama et al,].

After poly $(A)^+$RNA6μg, that was prepared first, was annealed with vector primer 2.2 μg first chain of cDNA was synthesized by carrying out the 1 hour reaction at37 deg C by adding reverse transcriptase (Seikagaku Corporation make) of 144 units.

After extracting phenol and precipitating ethanol, reaction mixture is reacted in the presence of 2.5μMdCTP at 37 deg C for 3 0 minutes with addition of 15 units of terminal transfer enzyme, dC tail attachment was also carried out.

After extracting phenol and precipitating ethanol, reaction mixture was digested at 55deg C for two hours in the 50 units of BstXI (New England Bio laboratory Corporation) .

After extracting phenol and precipitating ethanol, reaction mixture is annealed, after adding 300 units of colon bacillus DNA ligase, at 12 deg. C self ligation reaction was conducted for a night.

RNA chain was substituted by DNA by adding dNTP (dATP, dCTP, dGTP, dTTP), 300 units of colon bacillus DNA ligase, 20 units of colon bacillus polymerase, 15 units of colon bacillus RnaseH and by keeping it for one hour at 12deg C and next keeping it at 22 deg C for one hour.

Genetic transformation of colon bacillus NM 522 (Pharmacia) was conducted by using cDNA synthesis reaction mixture.

Genetic transformation was according to Hanahan Method. [D.Hanaha n, Journal of Molecular Biology (0022 - 2836, JMOBAK) 166: 557 - 580 (1983)].

[0025]

A portion of base sequence analytical above mentioned cDNA library of the human cDNA library was sowed in 2xYT agar medium containing 100μg/ml ampicillin and cultured at 37 deg. C for a night.

After picking up the colony of choice and inoculating 2xYT culture containing 100μg/ml ampicillin in 2ml and culturing it for 2 hours at 37 deg. C, helper phase MK13K07 is infected and further cultured at 37 deg. C for a night.

By centrifugation of culture solution, by separating cell mass and supernatant, according to the conventional method one chain of phage DNA was isolated from the supernatant.

After single chain DNA has conducted a sequence reaction using M13 sequence primer that has fluorescent pigment and Taq polymerase (Applied Biosystems Corp. Kit), fluorescent DNA is applied on the sequencer and the base sequence of cDNA is determined.

Reaction conditions were in accordance with the protocol belonged to the kit.

The acquired base sequence was converted to the amino acid sequence of three frames and the amino acid sequence database was created.

[0026]

cDNA cloning

Resulting from the search of above-mentioned database, in the partial amino acid sequence of bovine P28 protein, it has revealed that protein that is coded with plasmid pHP10097containing clone HP10097 has high resemblance with this partial amino acid sequence.

The structure of this plasmid is shown in Figure 1.

When entire base sequence of cDNA insertion was determined, it had the structure which consists of 5 not translated regions of 22bp, open reading frame of 681bp, and 3 not translated regions (Sequence Number 2).

Open reading frame codes protein comprised of 226 amino acid residue, as shown in table 1, in said protein, 4 partial amino acid of purified bovine P28 protein shown in the sequence number 3 to 6 and highly resembling amino acid sequence are included entirely..

[0027]

[Table 1]

Comparison of Table 1 amino acid sequence

********** & **********
****

Sequence Number amino acid sequence

(Position from N terminal) (1 character inscription)

********** & **********
****

1 (117 - 135) NRH EIA VML L EG GANPDAK					
				**********	
3				NRH EIA VML L EG GANPDAK	
1 (70 - 80)				DDAGWSPLHIA	
					*********
4				XDAGWQPLHIA	
1 (136 - 144) DHYEATAMH					
					*******
5				XHYEATAVH	
1 (190 - 204) LLVSQGASIYIENKEE					
					**********
6				XLVSQGASIYIENXEL	
*	: Section	that matches	with amino acid		Section

[0028]

Furthermore as a result of searching base sequence data base GenBank/EMB L/DDBJ making use of the sequence of cDNA that is acquired, it was understood that in EST database, the partial sequence of cDNA (record number R13947) that matches partially with cDNA of this invention shown in sequence number 2 is recorded.

However, because the partial sequence matches, it can not be assured that this fragment and the complete log cDNA of this invention are derived from the same mRNA.

In addition, amino acid sequence and functions of the protein that might be coded from this sequence only are not known.

[0029]

Protein synthesis with in-vitro translation

In-vitro translation was conducted with T_NT rabbit reticulocyte solution kit (Promega) using plasmid pH P10097 that possesses cDNA of this invention.

In this case [^{35}S] methionine was added; developed product was labeled with radio isotope.

Any reaction was conducted according to the protocol belonged to the kit.

After applying developed product on SD S-polyacrylamide-gel electrophoresis, autoradiography was conducted and molecular weight of translated product was sought.

As a result, cDNA of this invention generated a translated product of molecular weight approximately 26 kDa.

This value agrees within the experimental error with predicted molecular weight 24,427 of protein that is expected from base sequence, which is displayed in Sequence Number 2; it was shown that this cDNA is certainly coding the protein displayed in sequence number 2.

[0030]

Development with colon bacillus

After digesting plasmid pH P10097 1μg, with 20 units of Pvull and 20 units of Hind III, was applied on 0.08% agarose gel electrophoresis and approximately 700 bp DNA fragment was cut from the gel.

Next, expression vector pMK12 for E. coli which possesses SD arrangementand rrnBT1T2 terminator of tacpromoter、 meta pyro カテ car ゼ (Japan Unexamined Patent Publication Hei 2-182186 disclosure) 1;mu g digestion afterdoing, was applied on 0.8% agarose gel electrophoresis with PvuII of 20 unit and the HindIII of 20 unit, DNA fragment of approximately 3 kbp was cut from the gel.

After connecting both DNA fragments with ligation kit, genetic transfer of the colon bacillus JM109 was done.

Plasmid pMKP28-PvuII prepared from transformed host, recombinant, which is an objective with restriction enzyme cutting map, was verified.

[0031]

According to the protocol in attachment 2 oligonucleotide primer PR1 and PR2 (5';-GGGACGTC ATGGAGGGGTGT GTGTCT AA-3') (5';-GTC CAGCT GAGCATGCCCAGT GCAAT-3') were synthesized with automated DNA synthesizer (Applied Biosystems corporation)in accordance with protocol of attachment.

5' ends translated region of cDNA was amplified with PCR kit (Takara Shuzo Co. Ltd.) using plasmid pHP10097 1 ng and primer PR1, PR2 respectively 100pmole.

After phenol extraction, ethanol precipitation, digested with AatII of 20 unit (Toyobo Co. Ltd. (DB 69-053-8160))and with PvuII of 20 unit, reaction product was applied on 1.5% agarose gel electrophoresis and approximately 150bp of DNA fragment was cut and purified from the gel.

[0032]

After digesting plasmid pMKP28-PvuII 1μg with 20 units of AatII and 20 units of PvuII, it was applied on 1% agarose gel electrophoresis and 3.7kbp DNA fragment was cut from the gel.

With ligation kit this DNA fragment was connected with the latest 150bp DNA fragment prepared with PCR after that colon bacillus genetic transfer was done.

Plasmid pMKP28 was prepared from transformed host, recombinant which is an objective with restriction enzyme cutting map, was verified.

The structure of the acquired plasmid is shown in Figure 2.

[0033]

10ml of pMKP28/JM109 that was cultured for a night was suspended in 100 ml LB culture medium containing 100μg/ml of ampicillin, shook and cultured at 37 deg. C, when A_{600} has become 0.5 approximately, to make it 1mM isopropylthiogalactoside was added.

Furthermore after culturing it for 3 hours at 37 deg C, microbe collection was done centrifugally.

After ultrasonic treatment, when this solution was applied on SD S-polyacrylamide electrophoresis, a band of protein, which is induced to position of26 kDa, was recognized.

[0034]

[Effects of the Invention]

This invention provides human 26S proteasome P28, a DNA、 that codes said protein and a human cDNA that codes said protein.

Protein of this invention is useful in the elucidation of the detail functions of proteasome, and in the diagnosis and treatment of various diseases such as malignant tumor, where proteasome is involved.

In addition, said protein can be revealed in large scale by using DNA of this invention.

[0035]

Sequence Number: 1

Length of sequence: 678

Form of sequence: nucleic acid

Number of strands: double strand

Topology: straight chain

Kind of sequence: cDNA to mRNA

Sequence

ATG	GAG	GGG	TGT	GTG	TCT	AAC	CTA	ATG	GTC	TGC	AAC	CTG	GCC	TAC	AGC	48
Met	Glu	Gly	Cys	Val	Ser	Asn	Leu	Met	Val	Cys	Asn	Leu	Ala	Tyr	Ser	
1				5					10					15		
GGG	AAG	CTG	GAA	GAG	TTG	AAG	GAG	AGT	ATT	CTG	GCC	GAT	AAA	TCC	CTG	96
Gly	Lys	Leu	Glu	Glu	Leu	Lys	Glu	Ser	Ile	Leu	Ala	Asp	Lys	Ser	Leu	
			20					25					30			
GCT	ACT	AGA	ACT	GAC	CAG	GAC	AGC	AGA	ACT	GCA	TTG	CAC	TGG	GCA	TGC	144
Ala	Thr	Arg	Thr	Asp	Gln	Asp	Ser	Arg	Thr	Ala	Leu	His	Trp	Ala	Cys	
		35					40					45				
TCA	GCT	GGA	CAT	ACA	GAA	ATT	GTT	GAA	TTT	TTG	TTG	CAA	CTT	GGA	GTG	192
Ser	Ala	Gly	His	Thr	Glu	Ile	Val	Glu	Phe	Leu	Leu	Gln	Leu	Gly	Val	
	50					55					60					
CCA	GTG	AAT	GAT	AAA	GAC	GAT	GCA	GGT	TGG	TCT	CCT	CTT	CAT	ATT	GCG	240
Pro	Val	Asn	Asp	Lys	Asp	Asp	Ala	Gly	Trp	Ser	Pro	Leu	His	Ile	Ala	
65					70					75					80	
GCT	TCT	GCT	GGC	CGG	GAT	GAG	ATT	GTA	AAA	GCC	CTT	CTG	GGA	AAA	GGT	288
Ala	Ser	Ala	Gly	Arg	Asp	Glu	Ile	Val	Lys	Ala	Leu	Leu	Gly	Lys	Gly	
				85					90					95		
GCT	CAA	GTG	AAT	GCT	GTC	AAT	CAA	AAT	GGC	TGT	ACT	CCC	TTA	CAT	TAT	336
Ala	Gln	Val	Asn	Ala	Val	Asn	Gln	Asn	Gly	Cys	Thr	Pro	Leu	His	Tyr	
			100					105					110			
GCA	GCT	TCG	AAA	AAC	AGG	CAT	GAG	ATC	GCT	GTC	ATG	TTA	CTG	GAA	GGC	384
Ala	Ala	Ser	Lys	Asn	Arg	His	Glu	Ile	Ala	Val	Met	Leu	Leu	Glu	Gly	
		115					120					125				
GGG	GCT	AAT	CCA	GAT	GCT	AAG	GAC	CAT	TAT	GAG	GCT	ACA	GCA	ATG	CAC	432
Gly	Ala	Asn	Pro	Asp	Ala	Lys	Asp	His	Tyr	Glu	Ala	Thr	Ala	Met	His	
	130					135					140					

CGG	GCA	GCA	GCC	AAG	GGT	AAC	TTG	AAG	ATG	ATT	CAT	ATC	CTT	CTG	TAC	480
Arg	Ala	Ala	Ala	Lys	Gly	Asn	Leu	Lys	Met	Ile	His	Ile	Leu	Leu	Tyr	
145					150					155					160	
TAC	AAA	GCA	TCC	ACA	AAC	ATC	CAA	GAC	ACT	GAG	GGT	AAC	ACT	CCT	CTA	528
Tyr	Lys	Ala	Ser	Thr	Asn	Ile	Gln	Asp	Thr	Glu	Gly	Asn	Thr	Pro	Leu	
				165					170					175		
CAC	TTA	GCC	TGT	GAT	GAG	GAG	AGA	GTG	GAA	GAA	GCA	AAA	CTG	CTG	GTG	576
His	Leu	Ala	Cys	Asp	Glu	Glu	Arg	Val	Glu	Glu	Ala	Lys	Leu	Leu	Val	
			180					185					190			
TCC	CAA	GGA	GCA	AGT	ATT	TAC	ATT	GAG	AAT	AAA	GAA	GAA	AAG	ACA	CCC	624
Ser	Gln	Gly	Ala	Ser	Ile	Tyr	Ile	Glu	Asn	Lys	Glu	Glu	Lys	Thr	Pro	
		195					200					205				
CTG	CAA	GTG	GCC	AAA	GGT	GGC	CTG	GGT	TTA	ATA	CTC	AAG	AGA	ATG	GTG	672
Leu	Gln	Val	Ala	Lys	Gly	Gly	Leu	Gly	Leu	Ile	Leu	Lys	Arg	Met	Val	
	210					215					220					
GAA	GGT	678														
Glu	Gly															
225																

{0036}

Length of sequence: 1468

Form of sequence: nucleic acid

Number of strands: double strand

Topology: straight chain

Kind of sequence: From cDNA to mRNA origin: Name Of Specie: type of homo sapiens cell: phosphorus Ho ヲ cell line:U937 clone name: characteristic of HP10097 sequence: symbol expressing the characteristic: CDS existing position:23.703 method of determining characteristic: E

Sequence

AAGTAGTTGC	TGGGACAGCG	AA	ATG	GAG	GGG	TGT	GTG	TCT	AAC	CTA	ATG	GTC	52

							Met	Glu	Gly	Cys	Val	Ser	Asn	Leu	Met	Val
TGC	AAC	CTG	GCC	TAC	AGC	GGG	AAG	CTG	GAA	GAG	TTG	AAG	GAG	AGT	ATT	100
Cys	Asn	Leu	Ala	Tyr	Ser	Gly	Lys	Leu	Glu	Glu	Leu	Lys	Glu	Ser	Ile	
				15					20					25		
CTG	GCC	GAT	AAA	TCC	CTG	GCT	ACT	AGA	ACT	GAC	CAG	GAC	AGC	AGA	ACT	148
Leu	Ala	Asp	Lys	Ser	Leu	Ala	Thr	Arg	Thr	Asp	Gln	Asp	Ser	Arg	Thr	
			30					35					40			
GCA	TTG	CAC	TGG	GCA	TGC	TCA	GCT	GGA	CAT	ACA	GAA	ATT	GTT	GAA	TTT	196
Ala	Leu	His	Trp	Ala	Cys	Ser	Ala	Gly	His	Thr	Glu	Ile	Val	Glu	Phe	
		45					50					55				
TTG	TTG	CAA	CTT	GGA	GTG	CCA	GTG	AAT	GAT	AAA	GAC	GAT	GCA	GGT	TGG	244
Leu	Leu	Gln	Leu	Gly	Val	Pro	Val	Asn	Asp	Lys	Asp	Asp	Ala	Gly	Trp	
	60					65					70					
TCT	CCT	CTT	CAT	ATT	GCG	GCT	TCT	GCT	GGC	CGG	GAT	GAG	ATT	GTA	AAA	292
Ser	Pro	Leu	His	Ile	Ala	Ala	Ser	Ala	Gly	Arg	Asp	Glu	Ile	Val	Lys	
75					80					85					90	
GCC	CTT	CTG	GGA	AAA	GGT	GCT	CAA	GTG	AAT	GCT	GTC	AAT	CAA	AAT	GGC	340
Ala	Leu	Leu	Gly	Lys	Gly	Ala	Gln	Val	Asn	Ala	Val	Asn	Gln	Asn	Gly	
				95					100					105		
TGT	ACT	CCC	TTA	CAT	TAT	GCA	GCT	TCG	AAA	AAC	AGG	CAT	GAG	ATC	GCT	388
Cys	Thr	Pro	Leu	His	Tyr	Ala	Ala	Ser	Lys	Asn	Arg	His	Glu	Ile	Ala	
			110					115					120			
GTC	ATG	TTA	CTG	GAA	GGC	GGG	GCT	AAT	CCA	GAT	GCT	AAG	GAC	CAT	TAT	436
Val	Met	Leu	Leu	Glu	Gly	Gly	Ala	Asn	Pro	Asp	Ala	Lys	Asp	His	Tyr	
		125					130					135				
GAG	GCT	ACA	GCA	ATG	CAC	CGG	GCA	GCA	GCC	AAG	GGT	AAC	TTG	AAG	ATG	484
Glu	Ala	Thr	Ala	Met	His	Arg	Ala	Ala	Ala	Lys	Gly	Asn	Leu	Lys	Met	
	140					145					150					

ATT CAT ATC CTT CTG TAC TAC AAA GCA TCC ACA AAC ATC CAA GAC ACT 532
Ile His Ile Leu Leu Tyr Tyr Lys Ala Ser Thr Asn Ile Gln Asp Thr
155 160 165 170

GAG GGT AAC ACT CCT CTA CAC TTA GCC TGT GAT GAG GAG AGA GTG GAA 580
Glu Gly Asn Thr Pro Leu His Leu Ala Cys Asp Glu Glu Arg Val Glu
175 180 185

GAA GCA AAA CTG CTG GTG TCC CAA GGA GCA AGT ATT TAC ATT GAG AAT 628
Glu Ala Lys Leu Leu Val Ser Gln Gly Ala Ser Ile Tyr Ile Glu Asn
190 195 200

AAA GAA GAA AAG ACA CCC CTG CAA GTG GCC AAA GGT GGC CTG GGT TTA 676
Lys Glu Glu Lys Thr Pro Leu Gln Val Ala Lys Gly Gly Leu Gly Leu
205 210 215

ATA CTC AAG AGA ATG GTG GAA GGT TAAACAGCTT GGATTTATTC 720
Ile Leu Lys Arg Met Val Glu Gly
220 225

TTACTTTGTA TGTTGTGTTG TTGTCCCCAG TGTCCTACAA ACTAATGTAT TTGTGCACAA
GACATCATCT ATGAATGATG AAGTTTTCTC ACCTTCAAAG TCTTATAAAC ATGTTGACTC
TTGTTCCTGC TGAGTTACTT GTTCGAAGCT TACAGCTTGT TTTCCAGGCA TCGAATAACT
GTTGAGATTG TTCTACTGTT GTCGTATATT CTTCTATATT GAATTCTGGT TAATTTGGAG
TAACTAATTC TGTGGCTGTT GTGAGTCTTC AGCACCCTCC CATGTACCTT ATATCCCTCT
CTGAAACAGA ACAGCTCCAA TAGCAACAAG CTAGTTGTTC TGCCAGATGT TTCTATGTGG
ATTCTGTAAT GTTCCTCCAT ACAGTTAAAA CATCCTAACT TGTTTTTCAA GCTCACTCAG
GCCTACGCCA AACGTTTCTG TTTTTTTTAA CCATGAGGTT TAATTTATTT TTGTGATAGG
AGGGATATTT ACATATTTTA GTGGACCACA TTTTAAGTTG GATGGTGTGC TCTAAAATAC
TGAAAAACAA TAGCCCATAT ACCTATGTAT TTGTTTTTGA TGGGTTGTTT ACTCTGAAAT
AAAATGTATG GTTTTCTTAA AAGGAAGTTT TAAAGTACCT ATTTTGTGTC ATCCTGTATT
GAAAAGAATG TCAAGCTTGT TAAAATGACA TGTAACAAAA ATGTATTTTG ATTTGTATTT

CAGAAACTAA AAAATAAAAT GTTGAAAG 1468

{0037}

Length of sequence: 19

Form of sequence: amino acid

Topology: straight chain

Kind of sequence: kind: erythrocyte of peptide fragment type: intermediate section fragment origin: organism name: bovine cell

Sequence

His

Lys

{0038}

Length of sequence: 11

Form of sequence: amino acid

Topology: straight chain

Kind of sequence: kind: erythrocyte of peptide fragment type: intermediate section fragment origin: organism name: bovine cell

{0039}

Length of sequence: 9

Form of sequence: amino acid

Topology: straight chain

Kind of sequence: kind: erythrocyte of peptide fragment type: intermediate section fragment origin: organism name: bovine cell

{0040}

Length of sequence: 16

Form of sequence: amino acid

Topology: straight chain

Kind of sequence: kind: erythrocyte of peptide fragment type: intermediate section fragment origin: organism name: bovine cell

Sequence

Xaa	Leu	Val	Ser	Gln	Gly	Ala	Ser	Ile	Tyr	Ile	Glu	Asn	Xaa	Glu	Leu
1				5					10					15	

[Brief Explanation of the Drawing(s)]

[Figure 1]

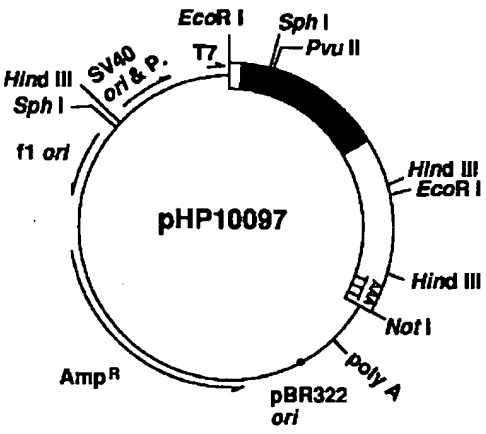
It is a figure that displays structure of clone HP10097.

[Figure 2]

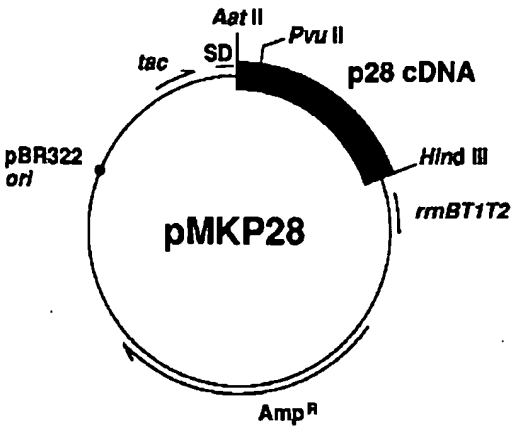
It is a figure that shows structure of developed vector pMKP28 for colon bacillus

Drawings

[Figure 1]



[Figure 2]


Aat II
Pvu II
SD
tac
p28 cDNA
pBR322
ori
Hind III
rrnBT1T2
pMKP28
Amp R